

NOV-30-2006 THU 12:40 PM

FAX NO.
RECEIVED
CENTRAL FAX CENTER

P. 15

Appl. No. 09/848,982
Amendment and Response

NOV 30 2006 Docket No. 85804.014508

REMARKS

Claims 1 to 43 are pending in the application, of which Claims 1, 12, 23 and 34, the independent claims, are being amended, and Claims 36 to 43 are being added. Reconsideration and further examination are respectfully requested.

Claims 1 to 35 are rejected under 35 U.S.C. § 103(a) over U.S. Patent No. 5,418,951 (Damashek) and U.S. Patent No. 6,272,456 (de Campos), and alternatively over Damashek and U.S. Publication No. 2002/0152204 (Ortega). Reconsideration and withdrawal of the claim rejection are respectfully requested.

Turning to the language of the claims, according to the method recited in Claim 1, a query string of characters is modified using a predetermined set of heuristics. A character-by-character comparison of the modified query string with at least one known string of characters in a corpus is performed in order to locate an exact match for the modified query string. In response to not finding an exact match for the modified query string, the method performs the steps of forming a plurality of substrings and using an information retrieval technique on the sub-strings in order to locate an equivalent for the modified query string. The plurality of sub-strings are formed from the query string by identifying one or more character sequences in said query string and a frequency of occurrence of each of the one or more character sequences in the query string, the formed sub-strings having varying lengths such that at least two of the formed sub-strings differ in length, on which sub-strings an information retrieval technique is used to identify a known string of characters equivalent to the query string.

Damashek and de Campos, are directed to a different problem than that of the claimed invention. More particularly, both Damashek and de Campos are directed to determining the language (or topic) of an unknown-language document using reference documents for which the

Appl. No. 09/848,982
Amendment and Response

Docket No. 85804.014508

language (or topic) is known. Identifying the language (or topic) of a unknown-language document is not the same as finding an exact match, or an equivalent in a case that an exact match is not located, for a query string. Furthermore, nothing in Damashek or de Campos, either alone or in any hypothetical combination (if such combination is even permissible, a point which is in no way conceded by Applicants), teaches, suggests or discloses the claimed two-fold approach for finding a match for a query string, which comprises performing a character-by-character comparison to locate an exact match, and if an exact match is not found forming sub-strings from the query string and using an information retrieval technique on the formed sub-strings to identify an equivalent for the query string.

Both Damashek and de Campos generate a score for a reference document based on the n-grams found in the unknown document which are also found in the reference document. If an n-gram found in the unknown document is not found in the reference document, both Damashek and de Campos simply generate a score for the reference document that reflects the fact that the n-gram was not found in both documents. Nothing in Damashek or de Campos teaches, suggests or discloses locating an equivalent of a query string by forming sub-strings from the query string and using an information retrieval technique on the formed sub-strings in response to an exact match for a query string not being found. While de Campos uses n-grams of different lengths, the n-grams are used to generate a score for a reference document that is then used to determine the language of an unknown-language document. Even assuming that the different-length n-grams correspond to the claimed sub-string (an assumption which is in no way conceded), nothing in de Campos teaches, suggests, or discloses use of different-length n-grams to locate an equivalent for a query string, let alone use of different-length n-grams to locate a known string

Appl. No. 09/848,982
Amendment and Response

Docket No. 85804.014508

that is an equivalent of a query string, and certainly nothing teaches, suggests or discloses formation or use different-length n-grams to locate a known string equivalent to a query string in response to a failure to locate an exact match for the query string.

Nothing in the applied art can be said to teach, suggest or disclose the claimed two-fold approach for finding a match for a query string, which comprises performing a character-by-character comparison to locate an exact match for the query string, and if an exact match is not found forming sub-strings from the query string and using an information retrieval technique on the formed sub-strings to identify an equivalent for the query string.

At col. 11, lines 7 to 35, de Campos describes a process for generating a reference document score by incrementally processing the unknown document. The process defines a fixed-length character window the size of an n-gram, and positions the character window over a portion of the unknown document to define the n-grams found in the unknown document. To score the reference document, the process then looks to see whether the n-gram is also one of the n-grams found in the reference document. The character window is then moved to define the next n-gram in the unknown document. More particularly, de Campos uses a three-character window and the string "ABCDE" to illustrate the process. The window is initially positioned at the beginning of the "ABCDE" string to identify the first 3-gram, "ABC". The "ABC" 3-gram is processed to determine a score for the reference documents and thereafter the character window is shifted one character to the right to identify the next 3-gram. Using this process, each one of the 3-grams found in the unknown document "ABC", "BCD", "CDE" are processed to generate a score for the reference documents. The incremental processing described in de Campos for scoring a reference document based on n-grams found in the unknown document is not the same

Appl. No. 09/848,982
Amendment and Response

Docket No. 85804.014508

as performing a character-by-character comparison of a query string with at least one known string in order to find an exact match for the query string. Furthermore, nothing in de Campos' incremental processing can be said to teach, suggest or disclose the claimed two-fold approach for finding a match for a query string, which comprises performing a character-by-character comparison to locate an exact match for the query string, and if an exact match is not found forming sub-strings from the query string and using an information retrieval technique on the formed sub-strings to identify an equivalent for the query string.

Nothing in Damashek or de Campos, either alone or in any hypothetical combination (if such combination is even permissible, a point which is in no way conceded by Applicants), teaches, suggests or discloses the claimed approach for finding a match for a query string, let alone the claimed two-fold approach for finding a match for a query string, which performs a character-by-character comparison to locate an exact match, and if an exact match is not found forms sub-strings from the query string and uses an information retrieval technique on the formed sub-strings to identify an equivalent for the query string.

Ortega has been reviewed and is not seen to remedy the above-noted deficiencies. Ortega describes a process of searching a database for information items using a query. While Ortega describes performing a character-by-character comparison, the character-by-character comparison is used to replace a misspelled query term with the correct spelling of the query term, i.e., to find a different query term as a replacement query term. Use of a character-by-character comparison to replace one query term with a different query term clearly cannot be said to be the same as performing a character-by-character comparison to find an exact match for the query string.

Appl. No. 09/848,982
Amendment and Response

Docket No. 85804.014508

Accordingly, Damashck, de Campos and Ortega, when taken alone or in any hypothetical combination suggested by the Office Action (if any such hypothetical combination is even permissible, a fact that is in no way conceded), are not seen to show each and every one of the claimed features.

For at least the foregoing reasons, Claim 1 is believed to be in condition for allowance. In addition, for at least the same reasons, Claims 12, 23 and 34 are seen to be in condition for allowance.

The other claims are each dependent from the independent claims discussed above and are therefore believed patentable for at least the same reasons. Because each dependent claim is also deemed to define an additional aspect of the invention, however, the individual consideration of each on its own merits is respectfully requested.

In addition to the elements of each of their respective independent claims, new Claims 37, 28, 40 and 42 each recite that the length of a sub-string is determined based on one or more character sequences identified in the modified query string and a frequency of occurrence corresponding to each one of the one or more character sequences, which is not disclosed in the applied art.

New Claims 36, 39, 41 and 43 have the added element of the weight for a given sub-string is based at least in part on a number of times the sub-string occurs in the query, which is not disclosed in the applied art.

In view of the foregoing, the entire application is believed to be in condition for allowance, and such action is respectfully requested at the Examiner's earliest convenience.

Appl. No. 09/848,982
Amendment and Response

Docket No. 85804.014508

The Applicant respectfully requests that a timely Notice of Allowance be issued in this case. Should matters remain which the Examiner believes could be resolved in a further telephone interview, the Examiner is requested to telephone the Applicant's undersigned attorney.

The Commissioner is hereby authorized to charge any required fee in connection with the submission of this paper, any additional fees which may be required, now or in the future, or credit any overpayment to Account No. 50-2638. Please ensure that the Attorney Docket Number is referred when charging any payments or credits for this case.

Respectfully submitted,



Carole A. Quinn
Reg. No. 39,000
Email: quinnc@gtlaw.com
Phone: (714) 708-6500

Date: November 30, 2006

Customer Number 32361
GREENBERG TRAURIG, LLP
Met Life Building
200 Park Avenue, 20th Floor
New York, New York 10166
Phone: (212) 801-9200
Fax: (212) 801-6400

OC 286107807